



# Conception d'un environnement informatique fondé sur la linguistique textuelle et destiné à la formation des enseignants de FLE en Colombie

Jorge Mauricio Molina Mejia

## ► To cite this version:

Jorge Mauricio Molina Mejia. Conception d'un environnement informatique fondé sur la linguistique textuelle et destiné à la formation des enseignants de FLE en Colombie. 5ièmes Rencontres des Jeunes Chercheurs en Environnements Informatiques pour l'Apprentissage Humain (RJC EIAH 2014), Université de la Rochelle, Jun 2014, La Rochelle, France. hal-01091862

**HAL Id: hal-01091862**

**<https://hal.science/hal-01091862>**

Submitted on 7 Dec 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Conception d'un environnement informatique fondé sur la linguistique textuelle et destiné à la formation des enseignants de FLE en Colombie

**Jorge Mauricio MOLINA MEJIA**

**4<sup>ème</sup> année de doctorat**

**Laboratoire LIDILEM, 1180 avenue centrale, BP25, 38040 Grenoble, France**

**Jorge-Mauricio.Molina-Mejia@e.u-grenoble3.fr**

## Résumé

L'objectif de nos travaux de recherche est de mettre en œuvre un système informatique d'aide à la formation des futurs enseignants de FLE en Colombie. Le système informatique que nous avons élaboré puise ses sources dans l'approche connue sous le nom de linguistique textuelle. Il possède deux caractéristiques essentielles : 1) il s'agit d'un dispositif de type hybride, c'est-à-dire, que d'un côté il pourra être utilisé pendant les cours de formation de manière présentielle et de l'autre côté, il pourra aussi servir d'exerciceur afin d'être employé en autonomie par les étudiants en formation ; 2) le système peut être paramétré afin que l'enseignant puisse créer ou éditer, lui-même, ses séquences pédagogiques.

## Introduction

Bien qu'utilisées depuis quelques décennies dans un but éducatif, les nouvelles technologies et notamment les TICE et les systèmes ALAO se sont focalisés dans l'enseignement et apprentissage des langues étrangères d'un public assez large d'apprenants. Peu ou très peu d'outils informatiques ont visé spécifiquement la formation des formateurs, c'est-à-dire, des futurs enseignants de langues, en particulier dans le domaine du FLE en Colombie (Molina & Antoniadis 2014). Cet « oubli » est à l'origine de notre réflexion et du système que nous présentons dans cet article.

Comme Chaudiron (2007) et Antoniadis (2010) le suggèrent, notre approche s'appuie sur les procédures et résultats du Traitement Automatique des Langues (TAL), seule possibilité actuellement pour tenir compte et exploiter les propriétés des langues Antoniadis (2010). Il tient ainsi

compte des besoins de ses utilisateurs (professeurs et futurs professeurs de FLE) et leur permet de manipuler des concepts familiers de leur domaine (Antoniadis 2010).

Un autre aspect à souligner est le fait qu'un système informatique doit être aisément utilisable et paramétrable par n'importe quel utilisateur n'ayant pas des connaissances spécifiques en informatique (Antoniadis 2010). Autrement dit nous avons conçu un outil informatique dans lequel, les composantes et les fonctionnalités pourront être paramétrées par l'enseignant lui-même, et il pourra créer ou éditer ses propres séquences pédagogiques grâce à des outils issus du TAL et de l'ALAO. Cette approche, celle qui permet de donner à l'enseignant l'opportunité de s'emparer de l'outil et de l'utiliser à ses fins, nous paraît très en concordance avec les pratiques d'enseignement qui prônent pour une plus grande liberté au moment d'utiliser des outils informatiques dans la classe.

## Genèse et contexte du projet

Notre projet est né de l'observation des étudiants en formation, futurs enseignants de FLE, qui faisaient leur stage professionnel à l'Université d'Antioquia en Colombie. À partir de ces observations, dans lesquelles nous avons constaté des problèmes au moment d'enseigner, nous avons décidé d'essayer de les résoudre en faisant appel à la linguistique textuelle. Ces problèmes étaient d'ordre de manque de connaissances au niveau linguistique, et se traduisaient dans le cours de français en une sorte d'insécurité linguistique.

À partir de ces premières observations, nous avons par la suite élargi le terrain de

recherche à d'autres universités colombiennes formatrices des futurs enseignants de FLE. Dans ces universités, nous avons réalisé des entretiens et au travers des questionnaires informatisés mis en ligne, nous avons cherché des données nous permettant de mieux comprendre notre problématique et de mieux cibler les besoins, ainsi que les actions à mettre en œuvre. C'est grâce à ces besoins que nous avons pu commencer à travailler tout d'abord à partir d'une approche méthodologique et ensuite à partir des activités leur permettant non seulement une meilleure préparation à l'enseignement mais également à être performants au niveau du discours linguistique et didactique.

Grâce à l'analyse des questionnaires de recherche nous avons pu comprendre que le problème est dû au manque d'une connaissance plus approfondie des notions linguistiques. Ce manque de connaissance se traduit par une sorte d'insécurité linguistique, autrement dit, étant donné qu'ils manquent de niveau de langue au moment d'enseigner, ils ont du mal à transmettre les connaissances à leurs apprenants.

### **Un projet fondé sur la linguistique textuelle**

La linguistique textuelle est une approche théorique relativement récente. Certains auteurs (Charaudeau & Maingueneau 2002) indiquent son émergence dans les années 1960. Ce qui est sans doute intéressant concernant cette approche théorique est le fait qu'au lieu de prendre la phrase comme unité d'analyse et de travail, elle utilise une unité beaucoup plus large, le texte (Adam 1990 ; 2011a). Ceci permet de faire des analyses de la langue à partir d'analyses du texte.

### **Notions en linguistique textuelle**

Parmi les notions qui font partie de la linguistique textuelle nous avons choisi trois et nous les avons divisées en d'autres sous-notions. L'idée est de faire découvrir aux étudiants en formation les différentes possibilités offertes par ce type d'approche linguistique.

Le choix des notions a été fait à partir de l'analyse des questionnaires de recherche et des réponses du public enquêté

(enseignants-formateurs, étudiants en formation et enseignants de FLE).

Les trois grands axes de travail sont :

- **L'analyse de la structure d'un texte :** dont l'analyse des phrases principales et secondaires, l'analyse de la structure logique d'un texte, et l'analyse du type et des séquences textuelles (Adam 2011b).

- **L'analyse de la cohérence et cohésion textuelles :** dont l'analyse de la coréférence textuelle (anaphores et cataphores) et l'emploi des connecteurs et marqueurs logico-temporels et argumentatifs.

- **L'analyse de la progression thématique :** dont les notions de thème et rhème.

C'est à partir de ces trois axes que nous comptons améliorer la performance linguistique des futurs enseignants de FLE. En ayant choisi de travailler à partir de l'écrit, et dans celui-ci, dans l'analyse du discours écrit, nous tenons à améliorer aussi les aspects liés à la langue orale. Car, en partant de l'écrit et des structures logiques, les étudiants en formation apprendront également à mieux s'exprimer tout en employant les bons éléments que ce soit au niveau de la cohérence et de la cohésion du discours (par exemple emploi des anaphores ou de cataphores afin de ne pas se répéter, ou la bonne utilisation des connecteurs, etc.).

### **L'intérêt d'un dispositif informatique fondé sur la linguistique textuelle**

En formation de formateurs, les outils informatiques de type ALAO conçus spécifiquement pour des étudiants en formation se font rares. Plus rares encore sont ceux qui prennent en compte la formation, en même temps, des aspects linguistiques et didactiques de la langue cible.

Certains dispositifs de formation, comme le « Français en première ligne » (Develotte & Mangenot 2010), par exemple, sont plutôt axés dans des échanges multimédias (communication médiatisée par ordinateur). Mais des systèmes informatiques qui cherchent à former à partir de la textualité sont presque inexistantes (Mangenot 1998).

C'est pourquoi, nous proposons notre système que nous allons décrire par la suite.

## Fonctionnalités et architecture

Nous avons conçu un environnement informatique qui est composé de deux interfaces (figure 1). Une première interface permet aux enseignants-formateurs de créer et de paramétrer des séquences pédagogiques selon leurs besoins pour la formation. Une fois créées, ces séquences sont envoyées aux étudiants en formation qui y auront accès à travers une autre interface qu'ils utiliseront et au travers de laquelle l'enseignant-formateur pourra faire un suivi du travail fait de leur part.

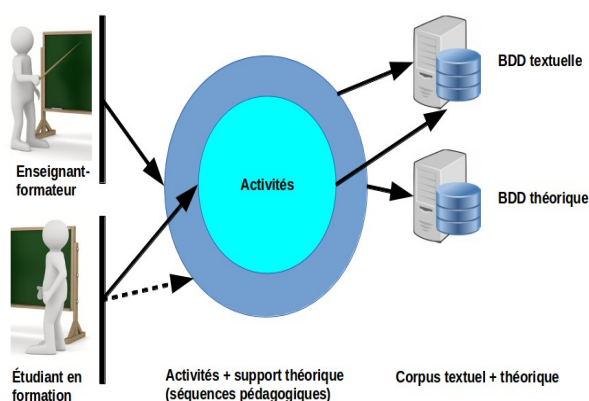


Figure 1. Architecture générale du système.

Pour ce faire, le système dispose de deux bases de données. La première est destinée au corpus textuel (BDD textuelle) et la deuxième contient le support et les schémas théoriques à être employés par les enseignants-formateurs, si besoin, pendant les séquences didactiques.

À partir de l'interface enseignant, les fonctionnalités permettent aux enseignants-formateurs de constituer des séquences de formation dites « à la carte ». Ce type de fonctionnalité permet à l'enseignant de préparer une activité pour un niveau déterminé, afin de travailler une notion en particulier et avec les textes qu'il considérera les plus adéquats (voir annexe A, modélisation des fonctionnalités du système informatique).

## Corpus textuel

### Type de textes

Nous avons constitué un petit corpus de textes de type journalistique et littéraire.

De petite taille car il compte à peu près 4000 mots, il vise les niveaux B1, B2 et C1 du CECR. Nous avons choisi des textes authentiques de la presse et de romans, qui ont été utilisés dans des anciens examens du DELF et du DALF. Ils apparaissent sur le site du CIEP.

L'idée est d'enrichir par la suite ce corpus avec plus de textes de type journalistique et littéraire et d'y ajouter des textes de vulgarisation scientifique, des extraits d'articles en didactique et en linguistique.

### Annotation des textes

Notre intérêt est de pouvoir exploiter de manière informatique les informations linguistiques portées par les textes. Pour ce faire nous faisons deux types d'annotations (voir figure 2). D'abord nous réalisons des annotations automatiques de type morphologique et syntaxique à l'aide de l'analyseur Cordial, version 20 (Synapse Développement 2013) ; une fois les annotations désambiguïsées nous procédons manuellement à une annotation fine des phénomènes liés à la linguistique textuelle. Le texte est structuré en format XML.

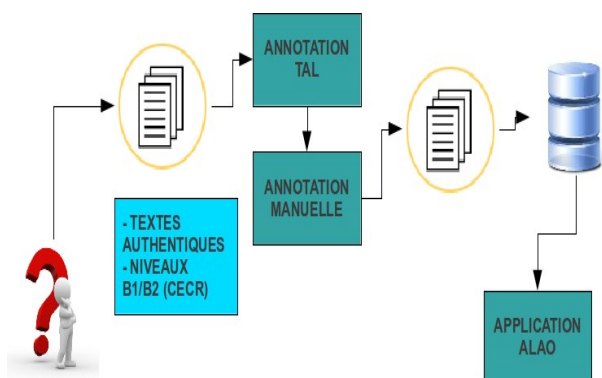


Figure 2. Schéma du traitement du corpus.

Dans la figure 3 nous pouvons apprécier une sortie de l'analyseur Cordial. Ces premières annotations nous les transformons dans un format XML.

La figure 4 montre un exemple d'étiquetage des éléments textuels, auxquels nous avons ajouté les informations linguistiques.

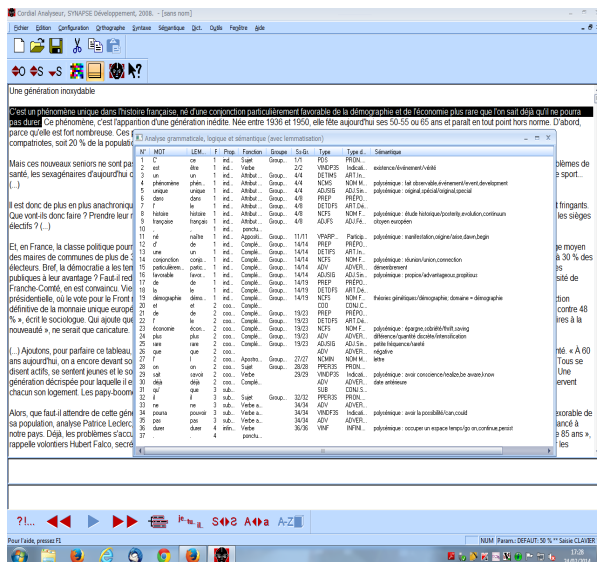


Figure 3. Exemple d'analyse du corpus à travers Cordial.

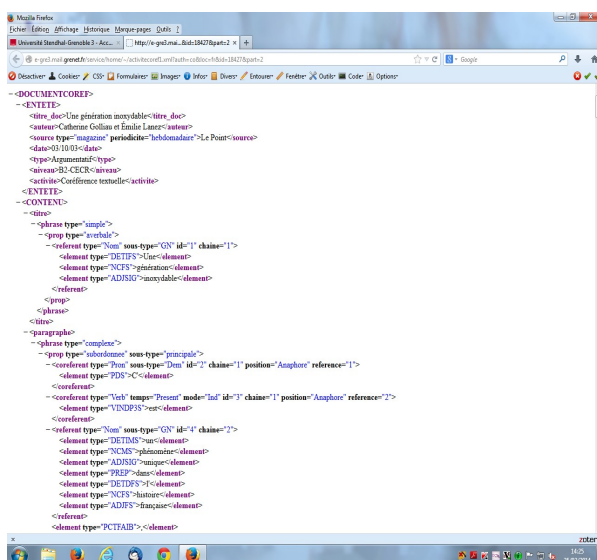


Figure 4. Exemple d'annotation en XML du corpus textuel.

Ce type d'annotation devra nous permettre la réalisation ultérieure d'activités.

## Scénarisation des activités pédagogiques

C'est en ayant annoté des textes et en mettant en relief des caractéristiques propres à des phénomènes linguistiques que nous voulons faire apprendre aux étudiants que nous pourrions créer des activités propres à la linguistique textuelle. Notre système informatique est conçu avec deux propriétés principales : la première est de faire travailler des notions

linguistiques ; la deuxième est de permettre la didactisation de ces notions.

## Séquences didactiques

Nous avons décidé de constituer notre système à partir des séquences didactiques, constituées de quatre phases (figure 5).

Comme nous pouvons le voir dans le schéma ci-dessous (figure 5), les phases 1, 3 et 4 utilisent la BDD textuelle dans laquelle nous avons stocké tous les textes qui ont été annotés avec les notions à travailler, tandis que la phase utilise le support théorique et les schémas d'enseignement seront stockés dans une deuxième BDD que nous avons nommée « théorique ».

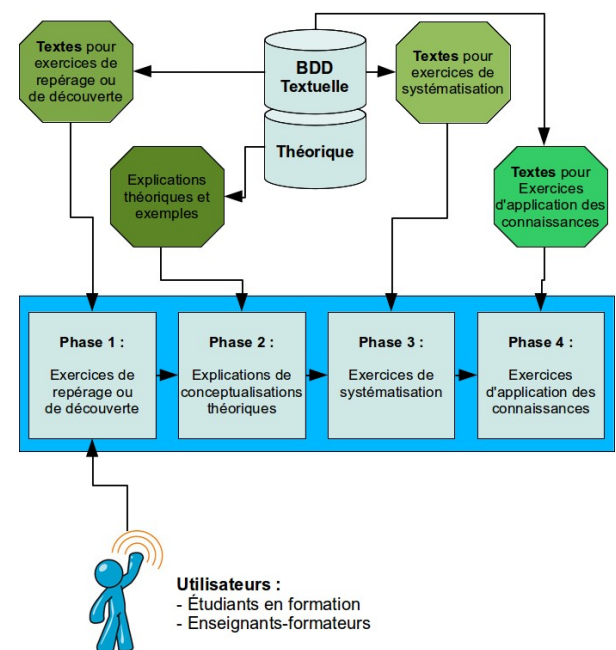


Figure 5. Schéma d'une séquence didactique.

Le rôle de chacune des quatre phases d'une séquence didactique est le suivant :

### Exercices de repérage ou de découverte (phase 1) :

Cette phase permet aux étudiants de découvrir un phénomène d'ordre linguistique en partant des activités qui permettent la déduction de ces notions.

### Explications de conceptualisations théoriques et didactiques (phase 2) :

Grâce à des explications théoriques d'ordre linguistique et des schémas explicatifs des notions liées à la linguistique textuelle, l'enseignant-formateur pourra les faire

acquérir aux futurs enseignants de FLE. Une deuxième étape sera liée à des explications de type didactique de la manière dans laquelle ces notions pourraient être travaillées en cours de langue.

### **Exercices de systématisation (phase 3) :**

Ce type d'exercices devrait permettre, aux étudiants en formation, de mieux saisir les notions travaillées pendant les séances en présentiel. Il s'agit des exercices (QCM, remplissage, déplacement d'éléments, etc.) permettant de mettre en œuvre d'une manière assez simple ces notions, pour mieux les apprendre et les comprendre.

### **Exercices d'application de connaissances (phase 4) :**

Dans ce type d'activité, les étudiants vont pouvoir créer des textes par rapport à la notion travaillée. Ils devront, par la suite, concevoir une activité en didactisant le texte créé qui aura été corrigé auparavant par l'enseignant-formateur.

### **Description d'une séquence didactique**

Nous passons maintenant à décrire le développement d'une séquence de formation à partir d'un exemple (tableau 1).

**Notion :** Cohérence et cohésion textuelle, analyse des connecteurs et marqueurs logico-temporels et argumentatifs.

**Niveau :** B2 du CECR

<b>Phase 1 : Activité de repérage (étape 1)</b>
<b>Actions :</b> <ul style="list-style-type: none"> <li>- Les étudiants vont d'abord explorer le texte et repérer tous les éléments qui connectent des paragraphes, des phrases ou des propositions entre elles.</li> <li>- Ils vont glisser ces éléments dans une boîte qui contient tous les connecteurs et les marqueurs de tout type confondu.</li> </ul>
<b>Feed-back :</b> <ul style="list-style-type: none"> <li>- Si les éléments appartiennent à la catégorie connecteurs et marqueurs logico-temporels et discursifs, ces éléments se mettent dans une couleur donnée, autrement ils restent dans la couleur initiale.</li> <li>- Les apprenants peuvent glisser des éléments entre le texte et la boîte jusqu'au moment où tous les éléments auront été validés.</li> </ul>

- Un compteur donne l'information de combien d'éléments manquent encore à être validés.

### **Phase 1 : Activité de repérage (étape 2)**

#### **Actions :**

- Une fois validée l'étape précédente, dans cette deuxième étape les étudiants vont classer tous les éléments trouvés.
- Ils glissent tous les éléments repérés et les mettent dans des cases contenant le type de connecteur ou de marqueur.
- Il s'agit tout d'abord de les classer en logiques, temporels, argumentatifs, et ensuite selon un type plus précis : spatiaux, énumératifs, reformulation, etc.

#### **Feed-back :**

- Il y a une validation au fur et à mesure de la réalisation de l'activité. Lorsque chacun des éléments est glissé dans chacune des cases ils prennent une couleur particulière selon la case correspondante.
- Si l'élément ne correspond pas à la case il restera dans la couleur initiale.
- Un nouveau compteur annoncera la quantité d'éléments manquants par case.

### **Phase 2 : Explication théorique linguistique**

- Dans cette phase les étudiants en formation ont accès d'un côté à des explications d'ordre théorique, et de l'autre côté à des schémas explicatifs par rapport à la notion travaillée.

### **Phase 2 : Explication didactique**

- Ici, les étudiants ont accès à des fiches pédagogiques avec des exemples, leur montrant la manière dans laquelle cette notion pourrait être didactisée.

### **Phase 3. Activité de systématisation (étape 1)**

#### **Actions :**

- Les étudiants doivent remplir des espaces avec les connecteurs ou marqueurs manquants. Que ce soit à partir d'exercices de type QCM, à menu déroulant ou à choix multiple, il s'agit de trouver l'élément qui convient selon le contexte.

#### **Feed-back :**

- Lorsque tous les éléments auront été choisis les étudiants cliquent sur *envoyer* et un feed-back montrant les éléments qui ne sont pas corrects s'affiche automatiquement.
- Les étudiants peuvent corriger plusieurs fois jusqu'à ce que tous les éléments soient complètement validés.



<b>Phase 3. Activité de systématisation (étape 2)</b>
<b>Actions :</b> - Une fois tous les éléments de l'étape précédente validés, les étudiants procèdent à les classer, grâce à des étiquettes qu'ils pourront choisir à partir des menus déroulants, choix multiples, etc.
<b>Feed-back :</b> - Comme dans l'étape précédente, il y a une validation à la fin de l'activité et la possibilité de corriger plusieurs fois.
<b>Phase 4. Activité d'application de connaissances (étape 1)</b>
<b>Actions :</b> - À partir d'une consigne donnée par l'enseignant, les étudiants vont écrire un texte dans lequel ils vont employer toutes les connaissances acquises. Dans ce cas précis, ils utilisent les connecteurs et marqueurs logico-temporels et discursifs. - Ils envoient le texte à travers leur interface informatique.
<b>Feed-back :</b> - L'enseignant-formateur est la personne chargée de corriger le texte et de l'envoyer aux étudiants en utilisant pour cela l'interface étudiant.
<b>Phase 4. Activité d'application de connaissances (étape 2)</b>
<b>Actions :</b> - Dans cette dernière étape, les étudiants en formation vont construire une séquence didactique en utilisant le texte qu'ils ont créé dans l'étape précédente. - Ils l'envoient par la suite au formateur en utilisant le même mécanisme qu'ils ont employé pour la première étape de la phase 4.
<b>Feed-back :</b> - Une fois l'activité corrigée, l'enseignant enverra l'évaluation aux étudiants à travers leur interface.

Tableau 1. Exemple de séquence didactique.

## Conclusion et perspectives

Nous sommes actuellement en cours de finition du processus de conception informatique de notre environnement d'aide à la formation. Il nous reste la mise en œuvre informatique et l'évaluation avant de le présenter dans les universités colombiennes concernées.

Nous sommes persuadés que cet outil informatique d'aide à la formation peut

s'avérer innovant pour la formation des futurs enseignants de FLE en Colombie. Il essaie de « tirer profit » de plusieurs champs disciplinaires (didactique de langues, linguistique textuelle, ingénierie pédagogique, TAL, ALAO, EIAH), et mettre leurs résultats au service des enseignants et futurs enseignants de FLE.

## Remerciements

Nous remercions ici notre directeur de thèse M. Georges ANTONIADIS et les membres du laboratoire LIDILEM de l'Université Stendhal Grenoble 3. Un très grand merci également aux universités colombiennes ayant participé au projet.

## Références

- Adam, J.-M. 1990. *Éléments de linguistique textuelle : Théorie et pratique de l'analyse textuelle*. Pierre Mardaga : Liège.
- Adam, J.-M. 2011a. *La linguistique textuelle*. Collection Linguistique Cursus, troisième édition. Armand Colin : Paris.
- Adam, J.-M. 2011b. *Les textes types et prototypes*. Collection Linguistique Cursus, troisième édition. Armand Colin : Paris.
- Antoniadis, G. 2010. « De l'apport pertinent du TAL pour les systèmes d'ALAO. L'exemple du projet MIRTO ». In *actes du 2e Congrès Mondial de Linguistique Française (CMLF-2010)*. 12-15 juillet 2010, La Nouvelle Orléans, USA.
- Chaudiron, S. 2007. « Technologies linguistiques et modes de représentation de l'information textuelle ». *Documentaliste-Sciences de l'information*, 2007/1 Vol. 44 pp 30-39.
- Charaudeau, P. & Maingueneau, D. 2002. *Dictionnaire d'analyse du discours*. Éditions du Seuil : Paris.
- Develotte, C. & Mangenot, F. 2010. *Former aux tutorats synchrone et asynchrone en langues*. Distances et savoirs vol. 8 - n°3/2010, p. 345-359.
- Mangenot, F. 1998. « Outils textuels pour l'apprentissage de l'écriture en L1 et en L2 ». Publié dans *Pratiques discursives et acquisition des langues étrangères* (Université de Franche-Comté, 1998), actes du 10e colloque international FOCAL, Besançon, 19-21 septembre 1996.
- Molina, J. M. & Antoniadis, G. 2014. « Toward the Constitution of a Hybrid Learning Environment for the FFL Teacher's Training in Colombian Universities Based on Text Linguistics ». In QUIROZ, Gabriel & PATIÑO, Pedro (Éditeurs) : *LSP in Colombia: Advances and Challenges*. Chapitre 15, Volume 175,

pp 233-249. Collection *Linguistic Insights*. Éditions Peter Lang : Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford, Wien.

## Annexe A : Modélisation système.

